DOCUMENT RESUME

ED 038 364 | 24 | SP 003 755

AUTHOR	Garrison, Jesse H.; Kersh, Bert Y.
TITLE	The Comparison of Two Forms of a Classroom Simulation Test Designed To Enhance Future Teachers Self-Definition and Teaching Style. Final Report.
INSTITUTION	Oregon Coll. of Education, Monmouth.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-9-I-033
PUB DATE	Oct 69
GRANT	OEG-9-94700330021
NOTE	27p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$1.45
DESCRIPTORS	Classroom Observation Techniques, Instructional Improvement, *Personality Tests, *Preservice Education, *Self Actualization, *Simulation, *Student Teachers, Supervisory Methods, Teaching Styles, Video Tape Recordings

ABSTRACT

The experiment reported in this document is based on the work of Garrison and others in developing a testing and interviewing procedure designed to enable increased self-definition on the part of prospective teachers and earlier development of appropriate and effective teaching styles. Three groups of elementary education majors were involved in the experiment, the immediate objective of which was to test the effectiveness of two different forms of the Classroom Simulation Test--a test based on motion picture sequences of classroom problems and students' responses to these. The control group interviews were conducted without any supporting test data. In both experimental groups a form of the Classroom Simulation Test (objective form in one group, projective form in the other) was used along with two other psychological tests (Edwards Personal Preference Schedule and the 16 Personality Factors test). The effectiveness of the interviews in promoting behavioral change was determined by judges' ratings of video tape recordings of each student's subsequent classroom performance. Results support earlier findings as to the effectiveness of the testing-interviewing technique and indicate: (1) There is a distinct advantage to using supporting test data in the interviews as contrasted with interviewing without the use of test data. (2) It may not matter which form of the Classroom Simulation Test (objective or projective) is used. (3) Learning effects are not limited to a single instructor personality. (JES)

ED038364

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Final Report

Project No. 9-I-033
Grant No. OEG994700330021

THE COMPARISON OF TWO FORMS OF A CLASSROOM SIMULATION TEST DESIGNED TO ENHANCE FUTURE TEACHERS' SELF-DEFINITION AND TEACHING STYLE

Jesse H. Garrison
Bert Y. Kersh

Oregon College of Education

Monmouth, Oregon 97361

October 1969

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

SP003755

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

	Page
Summary	1
Introduction	3
Self-definition	
Self-definition vs. phenomenological therapy	
Self-evaluation and self-direction	
Results of the 1966 pilot study	
Reasons for using the classroom simulation test	
Method	9
The experimental design	
Student population	
Orientation of students	
General testing and interviewing procedure	
Specific treatment using the objective test form	
Specific treatment using the projective test form	
Treatment of the no-test group	
Video-taping and rating each student's teaching behavior	
Results	15
Reliability and validity of the ratings	
Test of Hypothesis 1	
Test of Hypothesis 2	
Differences in instruction teams	
Conclusions and Recommendations	19
Conclusions	
Recommendations	
References	21
Appendixes	23

LIST OF TABLES

		Page
Table 1.	Mean ratings of teaching behavior for students in each group.	16
Table 2.	Results of t test of differences between mean ratings of teaching behavior for students in each group.	17

Summary

A pilot study conducted by Garrison in 1966 indicated that a testing and interviewing procedure with prospective teachers may be a very effective way to assist them in developing a more expressive and adaptive style of interacting with pupils in a classroom setting. Garrison attributed the students' improved performance to the fact that they gained a knowledge of their personal strengths and weaknesses (i.e., they "defined themselves") to a greater extent than students who did not participate in the interviews.

In 1967-68, Garrison developed a test called the "Classroom Simulation Test" which uses filmed classroom problems as test stimuli. Garrison reasoned that the classroom simulation test would prove to be a more effective medium for interviewing students than the standardized personality tests used in the 1966 study. The immediate objective of the experiment reported in the following pages was to test the effectiveness of two different forms of the classroom simulation test. Both test forms used motion picture sequences from the classroom simulation films as test stimuli. In the "objective test form," students responded to a prepared set of objective statements about each filmed problem after the films were shown. In the "projective test form," students wrote their responses freely to general questions ("What is the problem?", "How would you handle it?", etc.). The test data were interpreted to students in a series of private interviews by trained interviewers from two instructional teams. The effectiveness of the interviews was determined by judges' ratings of video-tape recordings of each student's classroom performance.

The results of the present experiment are consistent with earlier findings from the 1966 pilot study in which only standardized psychological tests were used. In the present experiment, the classroom test was used along with two other psychological tests. The data indicate that there is a distinct advantage to using supporting test data in the interviews as contrasted with interviewing without the use of test data. However, the evidence from the present study indicates that it may not matter which form of the classroom simulation test (i.e., objective or projective) is used.

If perfected, the testing and interviewing technique may greatly speed up the development of appropriate and effective teaching styles by student teachers. It is recommended that both basic and developmental research studies be conducted to answer critical questions concerning the psychological nature of "self-definition," and concerning the practical implications of using the testing and interviewing procedure on a broad scale.

Introduction

Pilot studies conducted by Garrison initially in 1966 (findings unpublished) indicated that the testing and interviewing procedure is highly effective. Students participating in the 1966 pilot studies completed a series of tests and interviews designed to assist them in "defining themselves." Afterwards, video-tapes of the teaching behavior of the participating students were compared with video-tapes of a control group of students who did not participate. Independent ratings by experienced teachers indicated that the experimental group of students was more "open" in their expression and reactions with pupils, and were more capable of adapting their instructional moves to the responses of the pupils.

All the tests employed in the 1966 studies were standardized psychological tests of personality which were difficult to interpret to the students in a way which would directly influence their teaching behavior. In an effort to develop an improved test, Garrison experimented with the use of classroom simulation films as test stimuli (unpublished pilot project, 1967-68, The Oregon CORD Project, USOE). In the 1967-68 project, a series of problems on film was selected from the original set of classroom simulation films developed by Kersh (1963, 1965). Two different forms of the classroom simulation test were developed: (1) an objective test in which the students accepted or rejected written statements about each film problem after viewing it; and (2) a projective test form in which the students wrote answers to general questions after viewing each film problem, e.g., "What is the problem?", "How would you respond?", etc. Both forms of the classroom simulation test may be administered in group settings, but the objective test form has obvious advantages in that it can be scored by machine techniques and evaluated objectively by use of test norms under development. The free-response (projective type) test form requires subjective scoring techniques and requires considerable skill on the part of the instructor in its interpretation. However, the free-response projective test has greater potential as a means for providing the future teacher with information about his strengths and weaknesses in interacting with students in a classroom setting. It also has the advantage of involving the student more than does the objective test because the instructional procedure relies on the student to define his own responses against normative data prior to the interpretive interviews with the instructor. The relevancy of the information provided the student through the free-response projective test reasonably would greatly enhance the effectiveness of the testing and interviewing instructional procedure, even though the test data would be of less value for screening students for acceptance or placement in teacher education programs.

Self-definition.

The value of self-definition, self-evaluation, and self-direction in the education of teachers is supported by both clinical and educational

sources. Jersild, Lazar and Brodkin (1962) support the value of this approach in their report that important effects on the professional work of teachers have been realized through fairly intensive individual treatment. Peck and Richek (1967), however, note that in the Jersild investigation there was a lack of quantitative data supporting the effect of therapy; and more importantly, that there was a lack of evidence on the question of whether the teachers' "improved" mental health did contribute to the promotion of positive mental health in their pupils.

We can gain additional insight into the value of an extensive investment in therapeutive facilities for teacher trainees by reviewing the reports of four programs included in the Association for Student Teaching, Forty-Sixth Yearbook (1967). The programs were sponsored by the National Institute of Mental Health and carried out at Bank Street College, San Francisco State College, the University of Wisconsin, and the University of Texas. The experience at the University of Texas provides a clear example. The Texas demonstration program began in 1958 with two major aims: "To produce teachers who are healthy and mature, mentally and emotionally, so that their effect on pupils will maximize good mental health; and to give these teachers a systematic education in the facts, principles, and practices of good mental health. The point of view that emerged from this was that "...counseling could be effectively employed to undergird professional education by producing deeper and full awareness on the part of the prospective teacher of his own personality and his probable interaction with the realities of the teaching role. At the same time we conclude that counseling per se was not the only, and probably far from the best, answer to this need. It cannot be relied upon as the exclusive vehicle for facilitating personal growth for all students...." (Peck, Bown and Veldman, p. 324-5).

In a concluding commentary on the National Institute of Mental Health (NIMH) studies, Fred Wilhelms noted that all four of the studies moved toward an emphasis "on 'the person inside the teacher'...what that person generally is." (AST, 46th Yearbook, p. 240)

Consider now an alternative to mental health therapy in enhancing the teaching effectiveness of teachers. The alternative approach is the one referred to in the introductory paragraphs above as an "instructional system of testing and interviewing." Individual treatment of students is involved, but the treatment is not therapy and it is not intensive treatment. Instead, standardized tests are used as tools providing data about the student teacher which are used as the basis for a series of private interpretive interviews between a student teacher and his college supervisor. The primary objective for the student is to gain an objective knowledge of himself (to arrive at a self-definition), then to assess his own strengths and weaknesses as they relate to classroom teaching situations, and to plan ways of adapting to changing classroom conditions.

Many teacher educators would agree that each potential teacher must have

some knowledge of, or means of discovering, the limits of his adaptability. We may assume a tremendous capacity for adaptation in our teacher trainees, but by college age each individual has built a self-structure that is extremely difficult to change. As a beginning in his efforts to develop an appropriate and effective style of teaching, the student of teaching must first discover those characteristics that he values most, i.e., those he sees as self-defining or as setting himself apart from others.

The definition of self does not on the face of it seem to constitute a major task, yet the peculiar situation that is encountered over and over in teacher education is that the student is unable to define himself in an objective way. The phenomenologists have demonstrated that, once an individual is aware of the difficulty of viewing himself objectively--a difficulty inherent in the placement of the sensory organs in the human structure--he is more willing and able to utilize available techniques for objectifying his perceptions. In fact, he will be able to invent means of "seeing himself."

There are a number of techniques from psychology and education that can be used as tools for the individual to obtain an objective definition of himself. A list of these would include assessment tests, both conventional and projective, individual consultation, particularly in the client-centered style of Rogers (1959), role playing, sensitivity training, interaction analysis (Amidon and Hough, 1967), micro-teaching (Stanford, 1967), and the repeated playback of performances recorded on audio and video tape (Jensen, 1966).

Self-definition vs. phenomenological therapy.

Millon (1967, page 2) points out that mental disorders (psychopathology) can be viewed from many different angles: (1) behaviorally, as complicated patterns of responses to environmental stress: (2) phenomenologically, as expressions of personal discomfort and anguish; (3) physiologically, as complex neural and chemical activity; and (4) intraphysically, as unconscious processes that defend against anxiety and conflict. From the standpoint of the teacher educator the phenomenological approach to therapy is probably the most practicable. For the classroom teacher, the behavioral approach is simply too cumbersome. The physiological approach requires a type of special education that classroom teachers cannot be expected to attain. The same is true of the intrapsychic approach. By contrast, the phenomenological orientation relies on behaviors which can be recognized by the maladjusted person, even when dealing with himself. The maladjusted person can even estimate the degree of his own disturbance by learning to react to his own feelings of personal discomfort and anguish relative to a particular activity in the classroom.

The instructional approach used in the present experiment is phenomenological in that it provides the student teacher with a method which enables him to reduce friction between himself and his environment, particularly the classroom setting. The instructional approach is not psychotherapy, however. Because they are carefully screened before being admitted to the teacher education program in the first place, it is likely that most student participants are more than usually capable of making adequate adjustments, particularly in a classroom setting.

Self-evaluation and self-direction.

A student's ability to evaluate himself and to decide on a course of action in the classroom is dependent largely on the extent to which he has defined himself, from the phenomenological point of view. Self-evaluation involves analysis of the effectiveness of one's self in a particular situation. Hopefully, most student teachers will know themselves well enough to be able to evaluate their potential effectiveness in the classroom before deciding to enter the teaching profession. After having arrived at a self-definition and having evaluated himself as having the potential for effective teaching, the student is then able to determine a course of action which will enable him to develop a teaching style that will make full and effective use of his unique personal resources.

Although a student teacher may arrive at a clear definition of himself very early in his college career, he will need a substantial knowledge of public education in America, the authoritative structure of a school system, limitations of school facilities, and the general duties and responsibilities of the teacher for his pupils--before he may be expected to evaluate himself effectively as a prospective teacher. Normally, the self-evaluation and self direction phase in the education of a teacher would come at about the junior year of college, i.e., after the student has had several introductory courses in professional education, including firsthand experience in a school classroom setting.

Results of the 1966 pilot study.

An initial exploration of the relationship of self-definition and teaching behavior was carried out by the principal investigator as part of a post-doctoral study program in 1966 and 1967 using psychological tests and interviews as the vehicle. Students taking part in the study were juniors and seniors enrolled in teacher education at Oregon College of Education. Of 80 students who volunteered to meet with the experimenter one hour per week for five or six weeks, 40 were selected to participate in the pilot study. Meetings were conducted with the students on an individual basis with all information treated as confidential and not shared in any way with the regular classroom instructors.

All students taking part were given the 16 Personality Factors test. The initial interview dealt with the student's response to the test and some effort was made to interpret the meanings of the various scores. All test results were shared with the student. Subsequent interviews dealt with an interpretation of the Edwards Personal Preference Schedule, GSR measures of student response to verbal stimuli, and a test of philosophical beliefs. The final input was a joint student-investigator analysis of a video-tape record of the student's teaching performance in the elementary classroom to which he was assigned.

The 40 student volunteers who participated in the study were assigned to experimental and control groups by a random process. The 20 Ss in the experimental group took the tests as described above and participated in between five and eight hours of individualized discussion time with the experimenter. Near the end of the second quarter of the study, all 40 students were asked to volunteer to be video-taped in their respective laboratory classroom settings. Due to time and scheduling difficulties, 19 video-tapes were actually made and eight of these were selected as representing the control group. The evaluation instrument (see Appendix B) was developed by the experimenter as a means of rating the performance of the students on the video-tape. Three independent judges, all qualified supervisors, were asked to view the tapes and rate each student on the variable of "use of expression," "reaction with pupils," and "style of presentation" as defined by the attached instrument. The following results can be reported:

1. The inter-rater reliability was extremely high. Specifically, the reliability of rater I to rater II was determined to be .987. Reliability of I to III was reported to be .761 and the reliability of II to III was reported at .925.

2. The statistical analysis yielded a significant difference at the .001 level in the mean ratings assigned to the experimental and control groups. The experimental group tended consistently toward the positive (lower numerals) end of the 10-point rating scale with the average ratings of the experimental group being 4.1 on the use of expression; 3.7 on reaction of pupils, and 4.3 on style of presentation. The control group produced 6.6 on use of expression, 7.1 on reactions of pupils, and 6.2 on style of presentation.

Students taking part in the study were interviewed during the progress of the two quarters. The subjective evaluations of the students were positive and demonstrated a high degree of enthusiasm. Most students were convinced that their basic feelings and attitudes were changing and, from their point of view, in a positive direction. Their reactions, on the most part, tended to support behavioral outcomes as manifested by the analysis of the video-tape.

Reasons for using the classroom simulation test.

One limitation of the testing procedure developed in the 1966 study was that the process of interacting with the student demanded a great deal of inferential reasoning. It often proved difficult for students to relate findings about themselves as revealed in a paper-pencil personality test to their anticipated behavior as a teacher in the classroom study. An investigation by Beaird and Schalock (1966) of motion picture films as test stimuli provided convincing evidence that a film test based on teachers' responses to filmed classroom episodes is a much more powerful predictor of teaching behavior than typical paper-pencil tests such as the Minnesota Teacher Attitude Inventory. It follows logically that asking a student teacher to predict how he would handle particular simulated (filmed) classroom problems might prove a much more effective way for him to arrive at an appropriate self-definition. An additional advantage of the classroom simulation test is that it would focus the student's attention more specifically on teaching instead of on his mental health status.

Method

The experiment was designed to test two hypotheses about the use of a classroom simulation test designed to help student teachers learn more about themselves. The classroom simulation test had been developed previously in connection with the 1967-68 project described in a previous section.

The experiment was designed to test the following two hypotheses:

> Hypothesis 1. Ratings of students who have been tested by either classroom simulation test form (and interviewed as described above) will be more positive than those who have been so interviewed without the use of test data.
>
> Hypothesis 2. Ratings of students who have been interviewed as described by using the projective form of the classroom simulation test will be more positive than those with whom the objective form is used.

The experimental design.

Originally, the design of the study called for 120 students rather than the 90 indicated in the scheme below. An additional control group had been scheduled which would not only have no tests administered, but also would have unstructured interviews instead of structured interviews as planned for the "no test" group referred to above. Unfortunately, however, one video-tape recorder failed to record either picture or sound for a large group of students in the classroom setting after the interviews were completed. The failure of the recording equipment was not discovered until the tapes were returned to campus and made available for assessment by the judges. It was not possible to retape the students without significantly altering the results, so the original design was collapsed by combining the original two control groups into the single "no test" group.

	Objective Test	Projective Test	No Test
Team A	n = 15	n = 15	n = 15
Team B	n = 15	n = 15	n = 15
	N = 30	N = 30	N = 30

The implications of the collapsed design are explained in greater detail in the sections which follow.

The experiment was designed to permit comparisons of the effectiveness of objective and projective test forms with each other and, in turn, with a control group interviewed without the use of test data. The design also permitted a separate comparison of two independent interviewers (members of Team A and Team B) across all treatments.

Student population.

All students taking part in the investigation were enrolled as juniors or seniors in elementary education at Oregon College of Education. They were enrolled in a two-quarter sequence of courses (called the "Junior Block") involving a study of educational psychology and educational methodology. The first quarter focused primarily on the content typically taught in the educational psychology and in reading methodology. This quarter included an assignment of approximately one-half day per week in a local elementary school classroom. The assignments for students included assessing their feelings toward children, their view toward the classroom in which they were visiting, and to some extent personal definitions of learning which seemed to be important to them in order to achieve their avowed goal of becoming a teacher. The formal teaching assignments were at a minimum during the first quarter and the emphasis was much more typically upon the feelings, perceptions, and attitudes which the student discovered in himself and in the learners with whom he was working. The second quarter concentrated more specifically upon teaching strategies and upon methodological approaches to assessment evaluation of teaching and of learning. The assignment in the classroom was extended to include one full day per week plus one full week during the quarter. The students were engaged more formally in the teaching act during the second quarter, especially for short periods of time devoted to teaching a single concept. The episodic teaching events were assessed both by the supervising teacher and by the student involved with the stress placed upon sensitive and open analysis of the student's performance rather than an emphasis upon proper or effective modes of instruction. The experiment took place at the time the students were involved in the second quarter sequence.

The 90 students were assigned by random processes to the three treatment groups. Instructors comprising the two teams were regular OCE faculty members normally assigned to teach the Junior Block course. They also taught the students educational psychology or general methods of teaching, and were actively involved with the participating students in other classroom experiences. In summary, the testing and interviewing system employed in the experiment was simply an extension of the actual course, and was considered by the student an actual assignment rather than an experiment for which they were volunteers.

Orientation of students.

The principal investigator was a member of the team of instructors designated to carry out the instruction in the two-quarter sequence as

well as to supervise the classroom experiences of the student enrolled. In this role he was able to appear before the students, explain the rationale and procedures of the project and ask that all students interested in taking part in the experiment volunteer to do so. From the list of volunteers, students were randomly assigned to the experimental groups. The initial explanation to the students emphasized that considerable faculty time was involved and therefore only a limited group of students would be able to take part. Those students who volunteered and were assigned to the "no-test" group were informed that their assignment occurred as a result of a random selection and therefore their role in the project would be limited.

General testing and interviewing procedure.

As was indicated in the previous paragraphs, there were two experimental groups which took one of two forms of the classroom simulation test, and a third group which took no tests at all. The following paragraphs describe the treatment which the test groups had in common. The specific treatment for the no-test group is described in a section which follows.

Prior to actual administration of the classroom simulation test, those students assigned to the objective and projective groups were given the 16 Personality Factors test and Edwards Preference Schedule. Profiles were constructed and individual interviews scheduled for each test profile. The interviews were typically scheduled for one-hour duration but with additional time made available for those students who for any reason wished to involve more time in the assessment of the tests and their reactions to them. There was a standard agreement throughout the experiment that the interviewers would attempt to have open office hours available so that the students taking part could make appointments or, in some manner, get in to see the interviewers whenever they felt a desire or inclination to do so.

The interviews as nearly as possible were conducted in the following manner:

a. The students and interviewers were jointly involved in exploring ideas and attempting to develop a system by which students might be helped to improve their teaching skills.

b. Both power and knowledge were shared jointly by the interviewer and the student. The interviewer neither approved or disapproved of any statements, concepts or ideas verbalized by the student.

c. There were no case studies made in the sense of tabulating the history of experiences of the students and then attempting to draw inferences to existing or future student behaviors, perceptions or beliefs.

d. The interviewer attempted to lead the student to identify his feelings and perceptions and to ask how these might be translated into useful teaching behavior. Little effort was made to define ways in which the student must "change" or "improve" his basic personality.

Students tended early in the interviews to treat their beliefs and their behavior largely as a function of the culture or the environment in which they had been reared. Their initial efforts were to interpret their behavior in terms of parental demands, early social and religious experience or certain inherited characteristics. Their approach to personal improvement reflected a direct and straightforward attack on what they saw as their personal and social limits. Statements such as, "I shouldn't be so hard on myself," or "I tend to be lazy," or "I worry too much," occurred quite frequently. The interviewer attempted to alter this line of statement into a different form by asking such questions as, "What kind of ethical or moral commitments do you now have which could make you vulnerable to feelings of inferiority or guilt," or "Do you feel your present goals are unrealistically high," or "Would life be better now if you could give up some of your fundamental commitments." The effort was to discuss the student's physical and intellectual capabilities, his religious and philophical commitments, and his personal and social outlooks in order to be as inclusive as possible when dealing with complex behaviors which a student otherwise might define quite narrowly, e.g., as a feeling of "inferiority." The assumption was that interpretations which tend to be honest, inclusive and broadly based represent a more hopeful approach to understanding one's behavior.

The intention of the interviewers was to remain neutral when listening to the ideas and value commitments of the students. However, the relationship of the student's personality profile and philosophical commitment was examined in each case at length. The purpose might be defined as an effort to assist each student in the process of establishing for himself a set of beliefs and commitments which were congruent with his makeup and, in the general sense, relevant to the profession of teaching. The specific content and form of each interview was governed insofar as possible by the manifest needs and purposes of the students involved.

Specific treatment using the objective test form.

The 30 students assigned to the objective treatment group took the classroom simulation test in a group setting (in addition to the two tests referred to above). After viewing each filmed episode they were

asked to respond to a set of objective statements about the problem on film (see appendix). The objective statements concerned the viewers' feelings and reactions toward each filmed problem, and possible solutions. The students were again scheduled for an interview in which the results were discussed. In the interviews they were asked to recapitulate the test episodes and their responses to them through the use of a series of slides taken from the films. The discussion in the interviews dealt with apparent similarities or differences in the classroom simulation test results and the results from personality tests.

Specific treatment using the projective test form.

The same steps outlined above were followed with the projective test group, with the following modifications. The students took the test in group situations and responded by writing answers to the questions dealing with their perception of the problem, their feelings about the problem, and their solutions or strategies to deal with the problem. In the subsequent interview, slides were used to remind the students of the situation and to ask them if they had further comments to make about them or if they wished to alter some of the comments they initially made. Their responses were then used as a basis for a discussion of the student's suppositions and concepts of himself as a teacher. Again there was no effort to get the student and the interviewer to agree but rather the effort was to get each to understand the other's point of view about the student as a teacher and to suggest areas where the student might need to re-examine, to increase effort or to attempt to be more aware of what was happening in the teaching role. Judgments about what the student ought to do were left as far as possible for the student to make in order to avoid a requirement that he comply with the interviewer's wishes.

Treatment of the no-test group.

Students in the no-test group were asked to examine and analyze their own experience background as a basis for the interviews. They were asked to come to the interviews with a prepared written statement of their own initial self-definition and, during the interviews, were asked to expand on their initial statement. They did not take either of the personality tests and did not view the classroom simulation films.

Video-taping and rating each student's teaching behavior.

The video-taping of each student's teaching behavior was completed during the second term of the two-term junior block sequence. As a regular part of the course experience, the students engaged in limited

practice teaching under supervision. The video-tapes were made of each student at approximately the same stage in their practicum, and the students were taped as they were carrying out equivalent laboratory teaching assignments. They were instructed to anticipate that they would be video-taped at some time, but were not told precisely when the taping would take place. The effort was to record a segment of each student's spontaneous teaching behavior. Each video-tape segment lasted approximately 20 to 30 minutes.

A rating procedure developed previously as part of the 1966 pilot study was used in this experiment (see rating form in Appendix B). The performance of each student teacher on video-tape was rated by three judges independently. A 10-point rating scale was used on each of three variables: (1) use of expression, (2) reactions with pupils, and (3) style of presentation. The numerical ratings per individual student made by each judge ranged from a high (positive) of 3 to a low of 30. Combining the ratings by the three judges resulted in composite scores for individual students ranging from 9 (high positive) to 90. The latter scores were used in the present experiment.

Results

Reliability and validity of the ratings.

During the 1966 pilot study coefficients of inter-rater reliability were recorded which ranged from .76 to .98 (three raters compared two-by-two). In the present experiment, analysis of variance was used to estimate test reliability. Rather than deriving a composite coefficient based on all 90 students, separate coefficients were calculated for each of the six experimental groups. The separate coefficients for each of the six subgroups ranged from .70 to .94 (see Appendix A for details).

The validity of the ratings is supported by the experience background of the judges. The three judges all were qualified supervisors who were well trained in the behavioral definitions of the three teaching variables rated ("use of expression," "reactions with pupils," and "style of presentation"). The rating form is included in the appendix.

Test of Hypothesis 1.

Table 1 shows the mean ratings of teaching behavior for students in each of the experimental groups and Table 2 shows the results of the statistical analysis.

According to the first hypothesis, the ratings of students who were tested by either classroom simulation test should be higher than those who were interviewed without the use of the simulation test data. In fact, they are, and the observed differences are statistically significant.

Test of Hypothesis 2.

The findings do not support the hypothesis that the projective test form is more effective than the objective test form. Clearly, with Team A it made little difference whether the objective or the projective form of simulation test was used. Both produced results which were superior to the no-test group. The results for Team B (although slightly contaminated) favor the use of the objective test form.

Table 1

Mean ratings[a] of teaching behavior for students in each group (n = 15 per group)

Instruction teams	Simulation test forms		
	Objective	Projective	No test
A	35.6	36.4	46.8
B	33.4	44.1	57.9

[a]Composite ratings by three judges, range 9-90. Positive teaching behavior indicated by lower ratings.

Table 2

Results of t test[a] of differences between mean ratings of teaching behavior for students in each group

	A-Obj	A-Proj	A-no	B-Obj	B-Proj
A-Obj	--	--	--	--	--
A-Proj	0.44	--	--	--	--
A-no	2.44	2.42	--	--	--
B-Obj	none	none	none	--	--
B-Proj	none	1.85	none	2.45	--
B-no	none	none	2.27	6.14	3.18

[a] $p<.05 = 1.70$ (28 df)
$p<.01 = 2.47$ (28 df)
"None" indicates that comparisons are not meaningful

Differences in instruction teams.

Since the pilot studies were based on the efforts of a single interviewer it is important to ascertain whether or not instructor differences markedly affect the differences attributed to test forms. The results indicate differences in the effectiveness of the interviewers and instruction teams A and B with respect to the particular form of the simulation tests used as the basis for interviews. However, the findings are consistent in showing the superiority of using some form of simulation test.

Conclusions and Recommendations

Conclusions.

The results of the experiment are consistent with earlier findings from the 1966 pilot study, at which time only standard psychological tests were used. In the present experiment, a classroom simulation test employing motion picture sequences of classroom problems was used along with two other psychological tests (Edwards Personal Preference Schedule, and the 16 Personality Factors test). The advantage of using supporting test data in the interviews was demonstrated by two independent teaching teams (Team A and Team B) indicating that the value of test data will remain for all instructors regardless of their styles of interviewing.

Although the data are not conclusive, they do indicate that it does not matter which form of the classroom simulation test (i.e., objective or projective) is used. In the present experiment, one interviewer (from Team A) evidently could have employed either form of the simulation test to equal advantage; whereas the other interviewer (from Team B) achieved better results by using the objective form.

Recommendations.

The experimental testing and interviewing procedure developed by Garrison is remarkably effective. The results of the 1966 study and current studies suggest that an extremely powerful instructional effect is in evidence. It is presumed at this point that after students have "defined themselves" with the help of appropriate test data and interpretive interviews with a skilled instructor, they almost immediately thereafter manifest behavioral changes in their expressive, reactive style with pupils in a classroom. Interviews alone are not sufficient; they should be based on test data. And the learning effects are not limited to a single instructor personality (the experimenter who developed the technique); evidently, others can accomplish the same desirable end. But there are yet many unanswered questions which should be researched. Following is a listing of questions suggested by the findings of completed research:

1. More precisely what characterizes the effective instructor-student interactions during the interviews which foster the student's understanding and acceptance of self?

2. Is the particular test medium called the "Classroom Simulation Test" really an important new dimension in the procedure of testing and interviewing? Or could almost any appropriate self-report inventory be employed by a

trained interviewer with equal effectiveness?

3. Precisely what is "self-definition," psychologically? Is it essentially a cognitive construct, or affective? Is the effect on teaching performance transient or stable? Can one's self-definition be altered?

4. Having defined himself, can the student change teaching styles more or less effectively?

Less psychologically important, but practical nonetheless, are questions concerning the cost of implementing such a testing and interviewing technique on a broad scale. Costs would reflect faculty development, changes in faculty load, and space requirements for a college whose resources are not initially so adaptive to the technique as OCE's.

It is strongly recommended that the line of developmental research indicated in the previous paragraphs be pursued systematically at OCE and elsewhere in the nation. Experiments of both the basic and developmental types are called for even in the few questions asked above. Findings would have far-reaching and fundamental implications, both psychologically and pedagogically.

References

Amidon, Edmund J. & Hough, John B. Interaction analysis: Theory, research and application. Reading, Massachusetts: Addison-Wesley. 1967, 402 p.

Benham, W. Gurney. Putnam's complete book of quotations - proverbs and household words. New York: G. P. Putnam's Sons, 1927, 1224 p.

Buros, Oscar K. Tests in print - A comprehensive bibliography of tests for use in Education, Psychology, and Industry. Highland Park, New Jersey: Gryphon Press. 1961, 479 p.

Bushnell, Don D. and Allen, Dwight W. (Eds.). The computer in American education, New York: Wiley. 1967, 300 p.

Jensen, Paul H. A study of self-evaluation applied to in-service teacher education. (Project proposal, Number 6-1470) Monmouth, Oregon: Teaching Research Division, Oregon State System of Higher Education. 1966.

Jersild, Arthur T., Lazar, Eve A. and Brodkin, Adele M. The meaning of psychotherapy in the teacher's life and work. New York: Teachers College, Columbia University. 1962, 151 p.

Kaplan, Louis Mental Health and Human Relations in Education. New York: Harper & Bros. 1959, 476 p.

Kersh, Bert Y. Classroom simulation: A new dimension in teacher education. Final report Title VII Project Number 886 - NDEA of 1958. 1963.

Kersh, Bert Y. Classroom simulation: Further studies on dimensions of realism. Final report Title VII Project Number 5-0848 - NDEA of 1958. 1965.

Millon, Theodore (ed.) Theories of psychology - Essays and critiques. Philadelphia: W. B. Saunders. 1967, 444 p.

Peck, Robert., Bown, Oliver H. and Veldman, Donald J. Mental Health in Teacher Education. Journal of Teacher Education. 1964, 15, (3, Sept.), 319-327 p.

Schalock, Henry D., Beaird, James H., and Simmons, Helen. Motion pictures as test stimuli: An application of new media to the prediction of complex behavior. Final Report Title VII Project Number 971, NDEA of 1958, 1964.

Stanford Teacher Education Program. Micro-teaching: A description. Palo Alto: Stanford University. 1967, 122 p.

The Association for Student Teaching. Mental health and teacher education. Forty-sixth yearbook. Cedar Falls, Iowa: The Association. 1967, 355 p.

Winer, B. J. Statistical principles in experimental design. New York: McGraw Hill, 1962.

APPENDIXES

Appendix A. Coefficients of Inter-rater Reliability for each Experimental Subgroup.

Appendix B. Affective Rating Scale.

NOTE: Detailed information about the Classroom Simulation Test may be obtained from Dr. Jesse H. Garrison, Chairman, Department of Education-Psychology, Oregon College of Education, Monmouth, Oregon, 97361.

APPENDIX A.

Coefficients of inter-rater reliability[a]
for each experimental group
(Three judges, n = 15 per group)

Instruction teams	Classroom Simulation Test Form		
	Objective	Projective	No-test
A	.88	.87	.80
B	.70	.94	.81

[a]By analysis of variance (Winer, 1962, p.131)

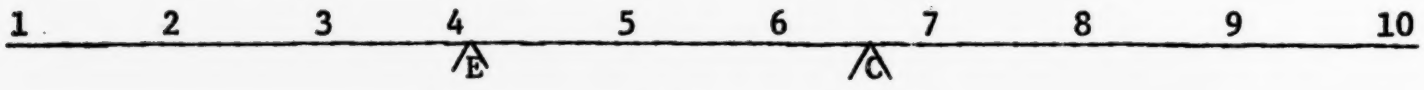
APPENDIX B.

Affective Rating Scale

USE OF EXPRESSION

dynamic voice qualities
facial expression smiles, frowns
bodily movements

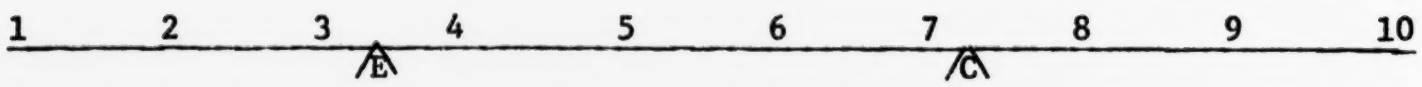
monotonous, dull voice
passive, fixed facial expression, stiff, stilted posture.



REACTIONS WITH PUPILS

responsive attentive
evidences interest and concern
elicits responses
accepts and recognizes students ideas

ignores, seems insensitive
interest only in "lesson"
gives questions/answers
stylized/dogmatic



STYLE OF PRESENTATION

flexible and open
exciting to students
interactive

rigid, stiff
boring, unreal to students
follows preconceived plan